CONTRIBUTI DI RICERCA CRENOS

UNIVERSITÀ DI CAGLIARI UNIVERSITÀ DI SASSARI
CENTRE FOR NORTH SOUTH ECONOMIC RESEARCH
CRENoS
CENTRO RICERCHE ECONOMICHE NORD SUD

MEASURING SPATIAL DISPERSION: AN EXPERIMENTAL TEST ON THE M-INDEX

Alberto Tidu
Stefano Usai
Frederick Guy

WORKING PAPERS

2022/06

arkadia

Centro Ricerche Economiche Nord Sud
(CRENoS)
Università di Cagliari
Università di Sassari

CRENOS was set up in 1993 with the purpose of organising the joint research effort of economists from the two Sardinian universities (Cagliari and Sassari) investigating dualism at the international and regional level. CRENoS' primary aim is to improve knowledge on the economic gap between areas and to provide useful information for policy intervention. Particular attention is paid to the role of institutions, technological progress and diffusion of innovation in the process of convergence or divergence between economic areas. To carry out its research, CRENoS collaborates with research centres and universities at both national and international level. The centre is also active in the field of scientific dissemination, organizing conferences and workshops along with other activities such as seminars and summer schools.
CRENoS creates and manages several databases of various socio-economic variables on Italy and Sardinia. At the local level, CRENoS promotes and participates to projects impacting on the most relevant issues in the Sardinian economy, such as tourism, environment, transports and macroeconomic forecasts.

www.crenos.unica.it
crenos@unica.it

CRENoS – Cagliari
Via San Giorgio 12, I-09124 Cagliari, Italia
tel. +39-070-6756397; fax +39-070- 6756402

CRENoS - Sassari
Via MURONI 25, I-07100 Sassari, Italia
tel. +39-079-213511

Title: MEASURING SPATIAL DISPERSION: AN EXPERIMENTAL TEST ON THE M-INDEX

Prima Edizione: Dicembre 2022

ISBN: 978 88 68514 389

Viale Bonaria 98 - 09125 Cagliari
Tel. 070/6848663 - info@arkadiaeditore.it
www.arkadiaeditore.it

Measuring spatial dispersion: an experimental test on the M-index

Alberto Tidu
University of Cagliari & CRENoS
Stefano Usai
University of Cagliari & CRENoS
Frederick Guy
Birkbeck College, University of London

Abstract:

In this paper, we assess the viability of a geographic approximation aimed to reduce the computational intensity necessary to measure spatial agglomeration with Marcon & Puech's (2017) M index. Indeed, despite representing a potentially very accurate way of measuring spatial distribution, M has not been sufficiently exploited so far because its computation needs crossing every point (i.e. firms, plants) with each other within the area under analysis: such a figure rapidly grows to unmanageable levels when the area is larger than a neighborhood or when every industry is taken into account. Consequently, practical applications of M have been exclusively experimental and circumscribed to very limited areas or to a handful of sectors. In our opinion, this is much regrettable since M provides many advantages compared to conventional measures of spatial distribution and also to other distance measures.

In order to verify whether a slight geographic approximation is tolerable – which would be consistent with Marcon & Puech's (2017, p. 30) assumption that "*cumulative functions are insensitive to errors at smaller scales than the distance they consider*" - we compute both actual M (with no approximation whatsoever) and approximate M for every industry in Sardinia. Our aim is to compare the results obtained when plants are located exactly where they are with those obtained when plants' positions are approximated to the centroid of the municipality where they are located.

We rely on a comprehensive dataset that allows us to identify the location, the specific industry and the number of employees for every single plant, and not only for firms as a whole. Our dataset's scope is not restricted to manufacturing, as it is often the case, but covers every area of activity, ranging from construction to transports and from retailers to other service industries. Moreover, we did not considered distance between approximated positions *as the crow flies*, but we relied on actual street distance and travel time between them: in the frequent case of orographically dishomogenous territories, it might be the case that such a measurement more accurately reflects the actual distance between establishment, than theoretical flying distance between actual locations.

If our approximation in the location of plants is positively outweighed by the great accuracy of M in operationalizing detailed geographic and economic information, then such an index could really be exploited for assessing agglomeration and dispersion patterns across space and along time, especially when much information is available, as it is ever more often the case.

Keywords: agglomeration; spatial methods; economic geography; distance-based measures; Sardinia
JEL Classification: R12; L60; L80

1 Introduction

Economic activity tends towards concentration and localization - both on a global and local scale -and it is therefore undeniable that studying agglomeration is essential to understand all sorts of economic phenomena and patterns of growth, in order *"to explain the riddle of uneven spatial development"* (Garretsen & Martin, 2010).

Historically, agglomeration studies had to rely on Gini, Theil or Herfindahl indices until the development of the so-called *dartboard approach* by Ellison & Glaeser (1997), which allowed to weight the actual distribution of activities against a theoretical distribution. However, even these measures were still biased by what is commonly referred to as the *Modifiable Areal Unit Problem* (Openshaw & Taylor, 1979). The Modifiable Areal Unit Problem affects all quantitative studies of spatial phenomena that rely on territorial aggregates - such as regions, provinces, municipalities, counties, etc. - as the unit of analysis. Indeed, when territorial aggregates are built on pre-defined borders, we have no way to distinguish spatial associations originating from the simple aggregation of data from real associations that do not depend on the type of spatial aggregation (Openshaw, 1984). A solution to the MAUP was presented by Duranton & Overman (2005) when they introduced *distance-based methods*: agglomeration indices whose measurement relies not on spatial aggregates but on the actual units, such as firms or plants. After Duranton & Overman's (2005) initial application, their innovative index *Kd* has been chosen by many scholars to measure agglomeration in a variety of countries all around the world, but the use of an absolute measure prevents results from being comparable across studies. Marcon & Puech (2010) provided a further improvement along this direction by developing a new cumulative measure, which they called *M* and which is able to account for both local and overall agglomeration, also allowing to assign a specific weight to each point. However, two significant drawbacks have limited its actual application so far: its computation intensity and the unavailability of data (Fratesi, 2004). The former is proportional to the squared number of points under review, resulting in extremely large figures when pairing every couple of plants in an entire country, or even within a single region. The other major difficulty - the availability of comprehensive micro-geographic data, needed to localize and assign reliable weights to each point - has been increasingly coped with in recent years, with the big data revolution providing researchers with a wide variety of data, often originating from unconventional sources as well (Piacentino, Arbia & Espa, 2021). It was indeed the availability of a comprehensive dataset – ISTAT's own ASIA – that led us to question whether there could be a less computationally intensive way of obtaining M for a decently sized area.

Our proposed solution is to verify whether M results show a high level of correlation with those obtained when approximating plants' locations to the centroids of the municipalities where they are located. In order to understand the difference in computational intensity that the viability of such an approximation would allow, it might be sufficient to highlight how the 9 billion pairs of plants necessary to compute M are reduced to about 10.000. We believe that – given the multitude of municipalities in Sardinia and their relatively small size – not much reliability would be lost through such an approximation. At first sight, such an approximation might seem counterintuitive when one is handling complex distance-based methods in order to pursue accuracy, but - as well expressed by Marcon & Puech (2017, p. 30) themselves -

"*cumulative functions are insensitive to errors at smaller scales than the distance they consider: if the uncertainty is a few hectometers, the number of neighbors up to a few kilometers is known with no error except for the more distant ones, which are a small proportion*". Moreover, our approximation shall not be perceived as a simple aggregation of economic data, since each plant is still considered and weighted separately from the others. Instead, what we do is approximating the geographic position of the plant by no more than a few kilometers[1]. We can also presume that most plants actually gravitate closer to the municipality centroid than a random distribution would predict, further reducing the magnitude of our approximation. Therefore, it is irrelevant that our methodology allows to simplify computations by numerically aggregating employees after their location has been registered, since – in a certain sense - this would occur even with the most pristine and precise implementation of distance methods: in the real world, employees are not piled up one above the other in the exact geo-localized position where the plant is registered, but they move around and are also separated from each other by at least a few meters, and – more often than not – much more than that, with many plants covering an ample surface (think of airports, harbors, large warehouses). Since nobody would require that a distance measure take into account vertical distances in a building or unavoidable separation between people in the same working area or even the *exact* position of each one of them in any given moment, we should only be concerned with one issue: whether the magnitude of our approximation is too large and whether it makes our results unreliable.

We believe that this is not the case, especially when considering the very high accuracy of the data we start with, compared to other similar studies: differently from much existing literature, we are distinguishing single establishments instead of relying on entire firms, not only as their geographic location is concerned, but also their specific industry and their number of employees; moreover, our dataset is not limited to manufacturing activities, but also covers services, which are usually excluded for a lack of reliable data. We believe that the loss in accuracy attributable to a slight geographic approximation is negligible when compared to the much more precise and comprehensive information available in respect to other studies.

Verifying whether correlation between *actual M* and *approximated M* holds has the side effect of providing us with interesting information about agglomeration and dispersion for every industry in Sardinia and their change during the critical years between 2007 and 2012: our preliminary review of the results and of their correspondence to literature and expectations will reinforce our trust in M's accuracy for describing agglomeration and dispersion.

Our paper begins with an overview of distance measures and their previous empirical applications within countries, regions and urban areas. We then proceed to describe our datasets and the methodology we followed for our analysis. Finally, we summarize our results and the major changes occurred between 2007 and 2012, showing the high correlation between results obtained with approximated plants' positions and results obtained with actual plants' positions.

---

[1] Only 2 municipalities - out of 377 - span more than 300 km² and the median surface is barely over 40 km².

2 Theoretical and empirical background

Evidence for a small group of European (Barlet, Briant & Crusson, 2013, Koh & Riedel, 2014), Asian (Nakajima, Saito & Uesugi, 2012) and American (Klier & McMillen, 2008 and Behrens & Bougna, 2015) countries confirms the widespread prediction that industrial activity exhibits specific location patterns. These findings suggest, therefore, that a high level of concentration in manufacturing can be observed in different countries of the world. Models describing and predicting agglomeration have been developed by economists as diverse as A.C. Pigou and Paul Krugman and such theoretical literature has been also accompanied by a sizeable amount of empirical studies aiming to measure agglomeration as accurately as possible (Tidu, 2021). These studies have relied on different *generations* (Nakajima, Saito & Uesugi, 2012) of indices before arriving to the current wave of distance-based methods.

The first generation corresponds to indices that rely on areal data to measure spatial concentration, such as Gini, Isard, Herfindahl, and Theil, where "*the precise location of firms is not available and the data only consists in aggregated counts over administrative zones*" (Bonneu & Thomas-Agnan, 2015, p. 291). In the study of Italy, three different first-generation measures were used by Pagnini (2003, p. 3) in order to measure agglomeration in manufacturing in 1996, showing "*that for an overwhelmingly majority of sectors centripetal forces prevail over centrifugal ones*". Other studies of concentration in Italy by means of first-generation indices were performed by Lafourcade & Mion (2007) and by De Dominicis, Arbia & De Groot (2013), with both providing a particular focus on the relationship between size and spatial agglomeration patterns.

The second generation started out when Ellison & Glaeser (1997) introduced the so-called *dartboard approach*, by way of comparing the degree of spatial concentration of employment in a given sector with the degree of concentration that would result if every plant in that sector were redistributed randomly across actually existing locations - that is, like darts thrown at the map. Ellison & Glaeser's index (henceforth, the *EG index*) would be used by Rosenthal & Strange (2001) to measure the level of spatial concentration among manufacturing industries at a 4-digit level for different geographic scales (zip code, county, and state) for the fourth quarter of 2000. Their aim was to explain differences in the spatial concentration of industries, by matching it with data on industry characteristics. Specifically, they regressed the EG index against a number of industry characteristics that they had identified as viable proxies for the three Marshallian forces of agglomeration – knowledge spillovers, labor market pooling, and input sharing - also controlling for product shipping costs and natural advantage. The EG index was also used by Kolko (2010), who was able to include services as well, without limiting his analysis to manufacturing. He also relied on a far deeper level of industrial detail, getting down to 6-digit industries. Studying US firms in 2004, he found that service industries, although more urbanized, were less agglomerated than manufacturing, possibly because transport costs represent an incentive to locate near their customers and because they are far less reliant on natural resources.

More recently, the necessity to deal with the so-called Modifiable Areal Unit Problem led to the development of a new assumption - continuous space – paving the road to a third generation of indices: *distance-based methods*. Proceeding from Ripley's (1976, 1977) seminal works and his K function, Duranton & Overman (2005) developed a new approach that allowed distance measures to be increasingly used to analyze spatial structures and agglomerations, without the need to rely on an approximation of space as discreet. Indeed - unlike more conventional measures ranging from Gini (1912) to Ellison & Glaeser (1997) indices – distance measures do not rely on any pre-defined zoning (i.e.: neighbourhoods, municipalities, communes, provinces, counties, regions), but on the distance between single points of interest, notwithstanding the geographical aggregation they – maybe only temporarily - belong to. Since they rely on the actual position of the target entities (such as individual plants or shops) and not on intermediary aggregates, distance-based methods can be a useful improvement compared to conventional spatial measures. Indeed, they are the only reliable way to overcome those issues that arise from referring to pre-defined zoning: geographic units are not necessarily homogenous, neither geographically nor economically, and therefore final values are dependent on the shape and size of the aggregation unit (since the distribution inside each area is lost through aggregation, and units at the opposite end of the same area are treated the same way as neighbouring units). Such an issue is commonly referred to as the *Modifiable Areal Unit Problem*[2] and its impact has been demonstrated, among others, by Kopczewska (2018). The distance measure introduced by Duranton & Overman (2005) is the following:

(1)
$$\widehat{K}_d(r) = \frac{1}{n\,(n-1)} \sum_{x_i \in R} \sum_{x_j \neq x_i,\, x_j \in \aleph} k\big(\|x_i - x_j\|, r\big)$$

where n denotes the total number of points, x_i are the reference points and x_j are its neighbors, with $k(\bullet)$ as a kernel estimator whose total sum is an estimate of the number of neighbors of x_i at the selected distance r

While researching distance-based methods, Duranton & Overman (2005) proposed five characteristics that sound distance measures should have:

1) It should be comparable across industries;
2) It should control for overall agglomeration trends across industries;
3) It should separate spatial concentration from industrial concentration;
4) It should be unbiased with respect to the degree of spatial aggregation;
5) It should provide an indication of the significance of the results.

A few years after its first introduction, Duranton & Overman's Kd was still the measure of choice when dealing with by then "booming" distance-based methods and the one that probably respected the largest number of properties listed above (Marcon & Puech, 2010). However, Marcon & Puech (2010) noted that most studies until then had not discussed an essential property of distance-based methods[3]: the difference between probability density

[2] Wong (2004, p. 572) notes that <<*Even though Gehlke and Biehl (1934) discovered certain aspects of the modifiable areal unit problem (MAUP), the term MAUP was not coined formally until Openshaw and Taylor (1979) evaluated systematically the variability of correlation values when different boundaries systems were used in the analysis*>>.

[3] With the exception of a short note by Duranton & Overman (2005) in the conclusion of their paper, where they argue that probability density functions reveal more information than cumulative functions do.

functions and cumulative functions. *Density functions* measure agglomeration *at* a specific distance from a reference point, whereas *cumulative functions* measure it *up to* a specific distance.

Marcon & Puech (2010) identify another dimension of distance-based methods: there can be *topographic*, *relative* or *absolute* measures, according to the reference value used to compare the distribution. A *topographic* reference uses physical space as a benchmark: the number of neighbors on a disk of radius *r* for a *cumulative function*, or on the ring at distance *r* for a *density function*. Topographic functions might simplify space - treating it as homogenous - or alternatively take into account the lack of homogeneity in the geographic space. A *relative* reference may use any other benchmark that is not physical space (e.g.: the distribution of plants that belong to every industry as a benchmark for the distribution of plants belonging to one specific industry). Finally, in the case of no reference, an *absolute* measure is defined, such as the absolute number of plants located at or within a given distance from a given one.

Marcon & Puech's *M* is a cumulative function that provides the relative frequency of neighbours of a given type (such as firms belonging to the same industry as opposed to the entire population of firms) within a certain distance, compared to the same frequency in the whole space. It is estimated by:

$$\text{(2)} \qquad \widehat{M}(r) = \sum_i \frac{\sum_{j \neq i} \mathbf{1}\left(\left\|x_i - x_j^c\right\| \leq r\right) w\left(x_j^c\right)}{\sum_{j \neq i} \mathbf{1}\left(\left\|x_i - x_j\right\| \leq r\right) w\left(x_j\right)} \Big/ \sum_i \frac{W_c - w(x_i)}{W - w(x_i)}$$

where x_j^c are neighbours of the chosen type, x_j are neighbours of any type, *r* is the selected distance, w is the weight of choice, W_c is the total weight of the first type of points, and $W$is the total weight of all points.

Recently, Kopczewska et al. (2019, p. 2412) developed a new measure of spatial agglomeration – the SPAG index – in order "*to determine to what extent the companies on the territory (e.g., in the region) are evenly distributed over space or follow the spatial agglomeration pattern*".

Duranton & Overman (2005) pioneered the application of distance-based measures for the study of agglomeration across industries in a developed country. They investigated location patterns in the manufacturing sector in the UK, by relying on their newly developed Kd index. They found that 52% of industries exhibited localization at a 5% confidence level, with 24% of them showing dispersion at the same confidence level, corresponding to a non-random distribution across space. This first contribution, which is both methodological and empirical, has been followed by many other studies which rely on their index to assess agglomeration levels across industries and, most importantly, their determinants along the line of Rosenthal & Strange (2001). Nakajima, Saito & Uesugi (2012) focused on Japan and found that about half of the 561 four-digit manufacturing industries they studied can be classified as localized, in contrast with a lower figure of only about 35% for service industries, also concluding that "*industries are becoming neither more concentrated nor more dispersed and the location patterns are stable over time*" (Nakajima, Saito & Uesugi, 2012, p. 18). Barlet, Briant & Crusson (2012) studied the location patterns of business-oriented service and manufacturing industries in France relying on an improved version of the Kd index, which takes into account the number of plants in each industry. They showed that concentration is more present among service industries (61%) than manufacturing industries (42%), especially at short distance. Researching Germany, Koh

& Riedel (2014) assessed the agglomeration patterns of four-digit industries in Germany using the Kd index. They found that 71% of manufacturing industries are localized while this ratio reaches 97% for the service industries. In line with the results above, Behrens & Bougna (2015, p. 48) found that "*depending on industry definitions and years, 40% to 60% of manufacturing industries are clustered*" and that localization in Canada has generally decreased during recent years. Cainelli, Ganau & Jiang (2020) demonstrated that different statistical techniques produce quite different pictures. In particular, they found that most Italian manufacturing industries experienced spatial dispersion processes during the period of the Great Recession. Finally, their results indicate that space–time dispersion processes occurred within small spatial distances and a short time horizon, although space–time interactions do not seem statistically significant.

As regards developing countries, the available evidence is scarcer, although some interesting contributions have recently appeared. Brakman, Garretsen & Zhao (2017) examined the location of manufacturing in China and found that around 80% of industries at 4-digit in China are significantly localized. Moreover, they found that localization increased rapidly in the period between 2002 and 2008, especially as a consequence of new entrants. Aleksandrova, Behrens & Kuznetsova (2020) analyzed the agglomeration and co-agglomeration patterns of manufacturing industries in Russia and found that 80% of 3-digit industries are both agglomerated and co-agglomerated. Almeida, Neto & Rocha (2020) found that almost 90% of Brazilian manufacturing showed statistically significant localization for 2006 and 2015.

Whereas applications of Duranton & Overman's Kd have been plenty, we have been unable to find a tentative measurement of agglomeration for every industry on a regional – or larger - scale through our measure of choice, Marcon & Puech's M. In order to find some empirical applications of *M*, one could turn to Jensen & Michel (2011) who used it to infer the spatial pattern of stores in Lyon (France), although this could be taken more like a mathematical exercise rather than an economic study[4]. Marcon & Puech (2015) themselves later developed such an application when "releasing" their newest measure, the lower-case *m*, in order to show how this could provide a different type of information in respect to Duranton & Overman's *Kd*[5], when describing the distribution of pharmacies in Lyon weighted against the distribution of non-food retail stores. Two other empirical applications of M were developed by Coll-Martìnez, Moreno-Monroy & Arauzo-Carod (2019) and Méndez-Ortega & Arauzo-Carod (2019) who, respectively, computed both *m* and *M* for creative industries and for software-developing industries in Barcelona metropolitan area, underlining how such measures provide the great advantage of being *relative* and not *absolute* (such as Duranton & Overman's *Kd*), thus comparable between industries and years. Also, Moreno-Monroy & Garcìa-Cruz (2016) used M to assess the degree of spatial agglomeration and co-agglomeration of *formal* versus *informal* manufacturing activity within Cali metropolitan area in Colombia. Finally, an interesting contribution has been provided by Zhang, Yao, Sila-Nowicka & Song (2021), who used both *M* and *m* to explore the geographic concentration of five manufacturing industries in the Chinese urban region of Jiangsu, relying on firm-level data. However, each one of the cited

---

[4] Points were not even weighted by the number of employees of the firm they were supposed to represent.

[5] It must be remembered that they both are *density measures*, not *cumulative measures* such as *M*.

contributions is limited either to individual industries and/or individual urban areas and is unable to study the whole economy of an entire region.

3 Data and methods

The exceptional level of detail provided by ISTAT's ASIA datasets in describing not only every firm, but every single plant in Italy, convinced us that we could still obtain reliable results, even if the geographical location of each establishment is approximated. Indeed, our aim is to show that such an approximation does not lead to questionable results and does not produce any significant loss as concerns accuracy.

As concerns the choice of Sardinia, we believe that dealing with an island prevents annoying edge effects that would make everything on the other side of the border disappear, strongly misrepresenting the actual economic activity of communities located on the outskirts, especially in the context of an open-border economy such as the European Union. However – except for Sicily and, indeed, Sardinia - every Italian island is way too small to have anything that even remotely resembles a real economy based on a wide array of industries: the largest minor island – Elba – consists of only 7 municipalities and its slightly over 30.000 inhabitants are disproportionately employed in touristic activities. On the other side, Sicily is the largest and by far the most populous Italian (and Mediterranean) island with over 5 million inhabitants, and such a demographic size signifies an extremely high number of pairs when every plant must be crossed with each other. Moreover, Sardinia – which is as large as Sicily and hosts a significantly lower population, but is still populated enough to sustain a sufficiently diversified economy – is also probably more akin to an autonomous economy because of its distance from the mainland, which is far larger than the tiny Messina Strait that separates Sicily from Calabria.

3.1 The dataset

ASIA (*Archivio statistico delle imprese attive*[6]) is a register established in 1996 in accordance with the provisions of European Council Regulation No. 2816/93 on Community coordination in drawing up business registers for statistical purposes, later replaced by Regulation (EC) No. 177/2008, and according to an harmonized methodology adopted by Eurostat.

Since 1996, ASIA covers every currently active enterprise[7] that contributes to gross domestic product, in the fields of manufacturing, trade and services, providing name, address, field of activity, number of employees, legal form, turnover class, and dates of creation and cessation. Economic activities not included in ASIA are: agriculture, forestry and fishing; public administration and defense; compulsory social security; activities of membership organizations; activities of households as employers; undifferentiated goods- and services-

[6] Italian for "*Statistical register of active enterprises*".

[7] Defined by ISTAT's quality report (https://www.istat.it/it/archivio/216767), in accordance with European Council Regulation No. 696/93, as "*the smallest combination of legal units that is an organizational unit producing goods or services, which benefits from a certain degree of autonomy in decision-making, especially for the allocation of its current resources. An enterprise carries out one or more activities at one or more locations. An enterprise may be a sole legal unit*".

producing activities of households for own use; activities of extraterritorial organizations and bodies; units classified as public institutions and private non-profit institutions. ASIA is updated every year through a process[8] that integrates several administrative and statistical sources[9], guaranteeing a proper statistical representation of active enterprises and of their identification, demographic and economic information. The register has a central role within economic statistics, and it is used for national accounting estimates.

Since 2004, ISTAT also provides another dataset, called Registro statistico delle Unità Locali (ASIA – UL), whose scope is roughly the same as the original register's and which has been built through a specific survey: Indagine sulle Unità Locali delle Grandi Imprese (IULGI). This survey has allowed to locate and define the main variables of each local unit[10].

3.2 The M index and its computation

Marcon & Puech (2010) noted how the largest number of properties identified by Duranton & Overman (2005) to define a sound distance measure, were respected by the latter's own measure, that is the *K-density* function (denoted *Kd*). However, since *Kd* is a density measure, they believed there was still the need for a cumulative function: indeed, the authors showed how the two types of functions are not substitutes, but complement each other, and, consequently, they created a new function named M[11], for the measurement of intra- and inter-industry geographic concentration.

In respect to a specific industry and for a selected distance range, five types of points – which, in our case, represent plants - are identified:

---

[8] ISTAT's quality report defines it as consisting in:

- Data acquisition;
- Analysis of the appropriateness of the sources;
- Transformation of data to standardize definitions;
- Transformation of data to standardize classifications;
- Record linkage;
- Audit and integration of unusual and/or missing data;
- Standardization, geocodification, de-duplication and validation of address data;
- Evaluation of consistency with previous data from the same elaboration.

[9] Agenzia delle Entrate; INAIL (Istituto Nazionale per l'Assicurazione contro gli Infortuni sul Lavoro); CCIAA (Camere di Commercio, Industria, Agricoltura e Artigianato); Banca d'Italia, INPS (Istituto Nazionale della Previdenza Sociale); Seat – pagine gialle Spa; ISVAP (Istituto per la Vigilanza sulle Assicurazioni Private e di Interesse Collettivo).

[10] Defined by ISTAT's quality report (http://siqual.istat.it/SIQual/visualizza.do?id=8889016), in accordance with European Council Regulation No. 696/93, as "*an enterprise or part thereof (e.g. a workshop, factory, warehouse, office, mine or depot) situated in a geographically identified place. At or from this place economic activity is carried out for which – save for certain exceptions – one or more persons work (even if only part-time) for one and the same enterprise*".

[11] Marcon & Puech (2010, pp. 747-748) "*called it the M function because it is an extension of the existing cumulative distance-based methods, namely Ripley's K function (1976, 1977) and Besag's L function (1977)*".

a) reference points (in our case, plants belonging to a specific industry);
b) target neighbor points (in our case, plants belonging to the same industry as and within the selected distance from a reference point).
c) target non-neighbor points (in our case, plants belonging to the same industry as the reference points, but outside the selected distance range from a reference point);
d) non-target neighbor points (in our case, plants that do not belong to the same industry as the reference points and are within the selected distance range from one of these);
e) non-target non-neighbor points (in our case, plants that do not belong to the same industry as the reference points and are outside the selected distance range from one of these).

The number of target neighbors is compared to both a local and a global benchmark: the former accounts for non-target neighbor points within the selected distance range, solving the potential issue that an industry might be "*more present in the Greater London than in Lincolnshire*" without being "*itself an indicator that such an industry is really clustered in London, where there are both more labour and more total manufacturing*" (Fratesi, 2004, p. 10); the latter, on the other hand, accounts for the relative abundance of said (neighbor and non-neighbor) target points in the entire area (i.e.: even outside the selected distance range) compared to (neighbor and non-neighbor) non-target points (in our case, plants that do not belong to the same industry that reference points belong to).

In other words, the average number of target neighbors is compared to a benchmark in order to verify whether they are more or less frequent than they would be if plants were distributed randomly. In order to control for the local density of points, target neighbor points (in our case, the number of plants, belonging to the same industry, located within the selected distance r from the reference point) are normalized by the number of all the neighbors located within the same radius. The average of the resulting ratio for each reference point will then be weighted against the same ratio for the entire area: if the former is higher than the latter – that is, M is greater than 1 - then the industry is somehow concentrated with points showing some degree of mutual attraction that would not be spotted if they were randomly distributed and independent from each other. On the other hand, if the latter ratio is higher than the former, it means that points tend to repel each other, therefore the industry is more dispersed than a random distribution.

3.3 Methodology

In our specific case, the equation

(3) $$\widehat{M}(r) = \sum_i \frac{\sum_{j\neq i} \mathbf{1}\left(\left\|x_i - x_j^c\right\| \leq r\right) w\,(x_j^c)}{\sum_{j\neq i} \mathbf{1}\left(\left\|x_i - x_j\right\| \leq r\right) w\,(x_j)} \Big/ \sum_i \frac{W_c - w(x_i)}{W - w(x_i)}$$

consists of:

a) reference points x_i corresponding to plants that belong to the industry for which the agglomeration index M is being computed;

b) target neighbor points x_j^c corresponding to plants neighboring the reference points and belonging to their same industry;
c) other "non-target" neighbor points x_j, corresponding to plants neighboring the reference points but not belonging to their same industry;
d) a distance range *r* - which takes the value of 5, 10, 15, 20 and 30 km *as the crow flies* – that "activates" the dummy when a plant x_j^c or x_j is within that distance from the plant x_i;
e) the number of employees *w* working in each plant;
f) the total number of employees W_c working in the industry for which the agglomeration index M is being computed;
g) the total number of employees *W* working in Sardinia.

The same computations are then repeated, substituting actual distance between plants with two different approximated distances, both provided by ISTAT through origin-destination matrices between Italian municipalities. Each plant's position is approximated to the centroid of the municipality where it is located, and computations are performed once again – this time using, respectively, road distance (in km) and travel time (in minutes) between such approximated locations. A similar expedient was used by Brakman, Garretsen & Zhao (2017) when studying spatial concentration of Chinese manufacturing firms: their limit was not computational, but concerned the actual location of the firms, since information was provided only at county level and they did not know exact addresses. They offered an interesting justification to their necessary approximation, by comparing the mean value of intra-county distances (19 kilometers) to the median value of all pair-wise distances between manufacturing firms in China (around 900 kilometers).

We are then able to check how much M results obtained through approximation of plants' locations correlate with M results obtained when actual plants' locations are taken into account. Since agglomeration is by definition a dynamic phenomen, we believe that our approximation needs to be validated not only for its precision in describing a static picture, but also for its accuracy in capturing the variations that occurr between different years. For this reason, we selected a period of particular turmoil: the Great Recession. We measured agglomeration for two different years[12]: the initial year is 2007, a year that ISTAT at the time described as "*exceptional as concerns firms' birth rate*"[13], showing a dynamicity that would not only be lost the following year, but probably was still unrecovered even a decade later. On the other hand, 2012 was the first year since the beginning of the Great Recession that showed an increase both in the number of firms and in the number of employees, although this would have later revealed itself as more of a rebound rather than a real recovery, since both firms and employees would then decrease every following year until 2016[14].

---

[12] It is not a coincidence that those same years were also chosen by Cainelli, Ganau & Jiang (2020), who acknowledged that 2007 "*is generally regarded as a pre-crisis year*" and that 2012 "*corresponds to the first year the Italian economy entered a second wave of downturn after the recovery peak reached in 2011*".
[13] *https://www.istat.it/it/files//2011/02/testointegrale20091006.pdf.*
[14] *https://www.istat.it/it/files/2018/12/C14.pdf.*

4 Baseline results

M results for the five distance ranges we computed are summarized in tables 1 and 2, where means and standard deviations are weighted by the number of plants in each industry. By construction, M can be computed only for industries with at least two plants, therefore those industries featuring only one plant[15] are not included in the results (although those plants were taken into account for the computation of other industries' M). Every industry featuring over five plants shows measurable agglomeration (i.e. it has at least two plants within less than 30 km from each other) and results are consistent between 2007 and 2012, showing strikingly similar means[16]. On the other hand, the apparently large difference ifrt-.t6n maximum values between 2007 and 2012 is entirely attributable to very small industries: if we only include industries with at least 10 plants and 100 employees, the largest value for M in 2007 amounts to 32,19 for *Manufacture of cement, lime and plaster (235)*, which is also the most agglomerated industry in 2012 with a remarkably similar value of 31,89.

Table 4 confirms our hypothesis that M does not lose much accuracy when it is computed after approximating plants' positions to the centroid of the municipality where each one is located[17]. Indeed, correlation between M computed with actual positions and M computed with approximated positions is extremely high, especially as concerns the 15- and 20-km distance ranges[18], and this is true when distance between municipalities is computed in travel time (minutes) or road distance (km). The high correlation between results obtained with approximate and with actual plants' positions also holds for changes in M values occurred between 2007 and 2012.

A comprehensive description is outside the scope of this paper, but even a general overview of the results might help to confirm that Marcon and Puech's M is able to measure accurately those agglomeration and dispersion phenomena that the existing literature describes and predicts.

As anticipated above, some of the industries that came out as the most agglomerated are extremely small in terms of both employees and plants, therefore we cannot put much trust in

---

[15] 8 in both 2007 and 2012.

[16] As expected, means are slightly higher than 1 and decrease towards such a value: this would be obtained by an industry whose plants were distributed in a pattern exactly mimicking the general distribution of every economic activity within the entire territory analyzed; since most industries tend to show some degree of agglomeration and since the distance ranges we selected are far shorter than a radius that could include the whole island, our values are easily explained.

[17] Industries are weighted by the number of plants they consist of, in order to account for large differences in their respective size.

[18] The lower correlations shown by shorter and longer distance ranges seem easily explainable: the former is probably affected by a proportionally larger impact of the approximation, whereas the latter is probably a result of orographic barriers which are not accounted for when distance is measured *as the crow flies*. Interestingly, if this is the case, it might be argued that our approximation does a better job than using actual distance between plants, since such barriers skew the result more than approximating the plant's position to a nearby point.

the significance of their results. However, some of those industries, despite their size, show consistently high M results for both years (tables 5 and 6), possibly hinting to the existence of actual agglomeration forces and not to mere coincidence: among them, *Manufacture of metal-forming machinery and machine tools (284)*, *Manufacture of tubes, pipes and hollow profiles and of tube or pipe fittings of cast-iron (242)* and *Manufacture of cutlery, hand tools and general hardware (257)*.

As concerns larger industries, *Manufacture of basic precious and other non-ferrous metals; reprocessing of nuclear fuels (244), Manufacture of knitted and crocheted apparel (143)*, *Manufacture of agricultural and forestry machinery (283)* and *Manufacture of refined petroleum products (192)* are consistently among the most agglomerated in both years.

When focusing on decently-sized industries, agglomeration is certainly not limited to manufacturing industries: indeed, *Camping grounds, recreational vehicle parks and trailer parks (553)* and *Sea and coastal water transport (501)* are consistently on top of the ranking in both years, as it is quite natural for activities respectively related to tourism and to water transport, as demonstrated – albeit less strongly – by *Hotels (551)* and *Other short term accommodation activities (559)*. Even more predictably, those industries "*for which location is constrained by natural advantages*" (Guillain & Le Gallo, 2010, p. 969) – such as *Building of ships and boats (301)* and, especially, *Quarrying of stone, sand and clay (081), Mining and quarrying n.e.c. (089)* and *Mining of hard coal (051)* – are all among the most agglomerated.

Whereas the persistency on the highest positions of the ranking might indicate actual agglomeration for small industries instead of mere chance, we might have a harder time when trying to infer dispersion for industries on the bottom: indeed, while we could argue that an industry consisting of five plants located close to each other for the best part of a decade might hint to an actual reason behind such proximity, the same number of plants located far from each other might very well not be driven by any particular dispersion force but by random chance. Therefore, it seems far more sensible to focus on those industries that manage to keep their plants decently far away from each other despite featuring many of them. Some of these industries are certainly predictable and this might be interpreted as a sign that our index is indeed representing dispersion as we would expect it to (tables 7 and 8):

- *Postal activities (531)*, *Monetary intermediation (641)*, *Waste collection (381)*, *Medical and dental practice activities (862)*, *Electric power generation, transmission and distribution (351), Other passenger land transport (493)* and *Wired telecommunications activities (611)* satisfy a public interest that requires them to be geographically dispersed and to follow the general population pattern rather than economic activity (with the latter supposedly more geographically concentrated than the former);
- *Electrical, plumbing and other construction installation activities (432)*, *Cleaning activities (812), Photographic activities (742)*, and most types of retailing activities, professional services and restoration do not directly satisfy a public interest *strictu sensu*, but still rely mostly on individual customers, with less room for economies of scale at plant level and, consequently, geographic concentration.

As concerns manufacturing activities, *Manufacture of structural metal products (251)* and *Manufacture of medical and dental instruments and supplies (325)* are the only decently sized industries that appear dispersed for every distance range both in 2007 and 2012.

Contrasting with Cainelli, Ganau & Jiang (2020, p. 443), who found that "*Italian manufacturing sectors experienced a process of space-time dispersion during the period of the Great Recession, although with slightly different intensity and patterns*", descriptive statistics provided in tables 1 and 2 show a tiny increase in the weighted mean of agglomeration results for every distance range. Indeed, even when focusing on the change for individual industries (accounting for the number of plants each industry consists of), there seems to be a slight percentage increase in M between 2007 and 2012, as summarized in table 3. Moreover, our findings also contrast with De Dominicis, Arbia & De Groot (2013, p. 5), who observed that "*whereas manufacturing has been spreading out, service activities have become increasingly clustered*", and instead they seem to suggest that manufacturing activities have generally clustered more than service industries, not less.

5 Conclusion

We have relied on comprehensive data provided by ISTAT – the Italian Institute of Statistics - to measure agglomeration for Sardinian industries in 2007 and in 2012. We believe our contribution is relevant with respect to both the methodological approach and the results obtained. Indeed, our operationalization validates an innovative way to use an accurate measure such as Marcon & Puech's M, whose experimentation had so far been restricted by its unmanageable computational intensity to the limited scope of individual city neighborhoods. This method, thus, extends its implementation possibilities to the study of larger geographic regions and even entire countries, as already pioneered by Tidu (2021). This is of the utmost importance because it offers an alternative to the passive acceptance of the distortions caused either by the Modifiable Areal Unit Problem or, alternatively, by the absence of a benchmark when relying on more commonly used distance-based methods, such as Duranton & Overman's *Kd*. With micro-geographic data becoming increasingly available (Arbia, 2001), it is crucial to learn how to exploit their whole potential when researching economics. Sardinia was chosen as the target of our study because of a demographic and economic size that make the island's data at the same time computationally manageable but economically relevant.

Our results show an extremely high degree of correlation between M computed with actual plants' locations and M computed by approximating the latter to the municipalities where they are located. This reinforces Marcon & Puech's (2017, p. 30) proposition that "*cumulative functions are insensitive to errors at smaller scales than the distance they consider: if the uncertainty is a few hectometers, the number of neighbors up to a few kilometers is known with no error except for the more distant ones, which are a small proportion*". Moreover, the results seem definetely plausible and in line with our expectations and with other researchers' findings, both in Italy and abroad. Indeed, when scrolling our ranking of the most agglomerated industries, it is easy to spot those factors that literature traditionally identifies as fundamental in generating agglomeration; and, on the other side of the spectrum as well, those industries that came out as the most disperse are certainly in line with literature predictions. Such results are surely interesting in and by themselves, but their relevance grows when they present the opportunity to assess the change that has occurred during such a dramatic event as the Great Recession. Specifically, we believe that some of the most at large considerations of previous literature were confirmed, with agglomeration somehow slightly decreasing (Behrens & Bougna, 2015; Almeida, Neto and Rocha, 2020) during the Great Recession, albeit with the most agglomerated industries – especially manufacturing ones – maintaining a high degree of agglomeration, and sometimes even showing an increase (Behrens & Bougna, 2015).

Although the tentative exploration of possible determinants is outside the scope of this work, our results point to the need of further study and interpretation about how agglomerations behave and react to the crisis. Especially, a significance test would provide us with the possibility to discern which industries actually produce reliable results, paving the road to an interesting exploration of possible determinants behind agglomeration and dispersion and to the identification of patterns for the development of a general model.

References

Aleksandrova, E., Behrens, K., & Kuznetsova, M. (2020). Manufacturing (co) agglomeration in a transition country: Evidence from Russia. *Journal of Regional Science*, *60*(1), 88-128.

Almeida, E. T., Neto, R. M. S., & Rocha, R. M. Manufacturing Location Patterns in Brazil. PhD dissertation in Economics. PIMES, Federal University of Pernambuco – UFPE.

Arbia, G. (2001). Modelling the geography of economic activities on a continuous space. *Papers in Regional Science*, *80*(4), 411-424.

Biagi, B., Dettori, B., Paci, R., & Usai, S. (2021). Economic development in Sardinia: overcoming the insularity gap. *Economic development in Sardinia: overcoming the insularity gap. Rivista Internazionale di Scienze Sociali, 2,* 165-186.

Barlet, M., Briant, A., & Crusson, L. (2013). Location patterns of service industries in France: A distance-based approach. *Regional Science and Urban Economics*, *43*(2), 338-351.

Behrens, K., & Bougna, T. (2015). An anatomy of the geographical concentration of Canadian manufacturing industries. *Regional Science and Urban Economics*, *51*, 47-69.

Biagi, B., Dettori, B., Paci, R., & Usai, S. (2021). Economic development in Sardinia: overcoming the insularity gap. *Economic development in Sardinia: overcoming the insularity gap.*, 165-186.

Bonneu, F., & Thomas-Agnan, C. (2015). Measuring and testing spatial mass concentration with micro-geographic data. *Spatial Economic Analysis*, *10*(3), 289-316.

Brakman, S., Garretsen, H., & Zhao, Z. (2017). Spatial concentration of manufacturing firms in China. *Papers in Regional Science*, *96*, S179-S205.

Cainelli, G., Ganau, R., & Jiang, Y. (2020). Detecting space–time agglomeration processes over the Great Recession using firm-level micro-geographic data. *Journal of Geographical Systems*, *22*(4), 419-445.

Coll-Martínez, E., Moreno-Monroy, A. I., & Arauzo-Carod, J. M. (2019). Agglomeration of creative industries: An intra-metropolitan analysis for Barcelona. *Papers in Regional Science*, *98*(1), 409-431.

De Dominicis, L., Arbia, G., & De Groot, H. L. (2013). Concentration of manufacturing and service sector activities in Italy: Accounting for spatial dependence and firm size distribution. *Regional Studies*, *47*(3), 405-418.

Duranton, G., & Overman, H. G. (2005). Testing for localization using micro-geographic data. *The Review of Economic Studies*, *72*(4), 1077-1106.

Ellison, G., & Glaeser, E. L. (1997). Geographic concentration in US manufacturing industries: a dartboard approach. *Journal of political economy*, *105*(5), 889-927.

Fratesi, U. (2004). *Measuring and explaining localisation: Evidence from two British sectors*. Università commerciale Luigi Bocconi, CERTet.

Garretsen, H., & Martin, R. (2010). Rethinking (new) economic geography models: taking geography and history more seriously. *Spatial Economic Analysis*, *5*(2), 127-160.

Gehlke, C. E., & Biehl, K. (1934). Certain effects of grouping upon the size of the correlation coefficient in census tract material. *Journal of the American Statistical Association*, *29*(185A), 169-170.

Gini, C. (1912). Variabilità e mutabilità. *Reprinted in Memorie di metodologica statistica (Ed. Pizetti E).*

Guillain, R., & Le Gallo, J. (2010). Agglomeration and dispersion of economic activities in and around Paris: an exploratory spatial data analysis. *Environment and Planning B: Planning and Design*, *37*(6), 961-981.

Jensen, P., & Michel, J. (2011). Measuring spatial dispersion: exact results on the variance of random spatial distributions. *The Annals of regional science*, *47*(1), 81-110.

Klier, T., & McMillen, D. P. (2008). Evolving agglomeration in the US auto supplier industry. *Journal of Regional Science*, *48*(1), 245-267.

Koh, H. J., & Riedel, N. (2014). Assessing the localization pattern of German manufacturing and service industries: a distance-based approach. *Regional Studies*, *48*(5), 823-843.

Kolko, J. (2010). Urbanization, agglomeration, and coagglomeration of service industries. In *Agglomeration economics* (pp. 151-180). University of Chicago Press.

Kopczewska, K. (2018). Cluster-based measures of regional concentration. Critical overview. *Spatial statistics*, *27*, 31-57.

Kopczewska, K., Churski, P., Ochojski, A., & Polko, A. (2019). SPAG: Index of spatial agglomeration. *Papers in Regional Science*, *98*(6), 2391-2424.

Lafourcade, M., & Mion, G. (2007). Concentration, agglomeration and the size of plants. *Regional science and urban economics*, *37*(1), 46-68.

Lang, G., Marcon, E., & Puech, F. (2020). Distance-based measures of spatial concentration: introducing a relative density function. *The Annals of Regional Science*, *64*(2), 243-265.

Marcon, E., & Puech, F. (2010). Measures of the geographic concentration of industries: improving distance-based methods. *Journal of Economic Geography*, *10*(5), 745-762.

Marcon, E., & Puech, F. (2017). A typology of distance-based measures of spatial concentration. *Regional Science and Urban Economics*, *62*, 56-67.

Méndez-Ortega, C., & Arauzo-Carod, J. M. (2019). Locating software, video game, and editing electronics firms: using microgeographic data to study Barcelona. *Journal of Urban Technology*, *26*(3), 81-109.

Moreno-Monroy, A. I., & Cruz, G. A. G. (2016). Intra-metropolitan agglomeration of formal and informal manufacturing activity: Evidence from Cali, Colombia. *Tijdschrift voor economische en sociale geografie*, *107*(4), 389-406.

Nakajima, K., Saito, Y. U., & Uesugi, I. (2012). Measuring economic localization: Evidence from Japanese firm-level data. *Journal of the Japanese and International Economies*, *26*(2), 201-220.

Openshaw, S. (1984). Ecological fallacies and the analysis of areal census data. *Environment and planning A*, *16*(1), 17-31.

Openshaw, S., & Taylor, P. (1979). A million or so correlation coefficients: three experiments on the modifiable unit area problem. *Statistical Applications in the Spatial Sciences*, 127-144.

Pagnini, M. (2003). Misura e determinanti dell'agglomerazione spaziale nei comparti industriali in Italia. *Rivista di Politica Economica*, *93*(2), 149-149.

Penttinen, A., Stoyan, D., & Henttonen, H. M. (1992). Marked point processes in forest statistics. *Forest science*, *38*(4), 806-824.

Piacentino, D., Arbia, G., & Espa, G. (2021). Advances in spatial economic data analysis: methods and applications. *Spatial Economic Analysis*, *16*(2), 121-125.

Ripley, B. D. (1976). The second-order analysis of stationary point processes. *Journal of applied probability*, *13*(2), 255-266.

Ripley, B. D. (1977). Modelling spatial patterns. *Journal of the Royal Statistical Society: Series B (Methodological)*, *39*(2), 172-192.

Rosenthal, S. S., & Strange, W. C. (2001). The determinants of agglomeration. *Journal of urban economics*, *50*(2), 191-229.

Sideri, M., & Usai, S. (2016). Economic Outlook: A missed opportunity. In *Surrounded by Water, Landscapes, Seascapes and Cityscapes of Sardinia (A. Cordale and G. Sistu eds.)*. Cambridge Scholars Publishing.

Tidu, A. (2021). Economic agglomeration in Italy before and after the Great Recession. *PhD Dissertation*.

Zhang, X., Yao, J., Sila-Nowicka, K., & Song, C. (2021). Geographic concentration of industries in Jiangsu, China: a spatial point pattern analysis using micro-geographic data. *The Annals of Regional Science*, *66*(2), 439-461.

Tables and figures

Table 1 – *Statistics for actual M results - Sardinia (2007)*

Variable	Obs	Mean	Std. Dev.	Min	Max
M (5 km)	224	1,43	2,15	0	261,90
M (10 km)	224	1,25	1,25	0	236,82
M (15 km)	224	1,17	0,67	0	101,50
M (20 km)	224	1,13	0,52	0	72,65
M (30 km)	224	1,07	0,24	0	16,87

Source: Compiled by the authors

Table 2 – *Statistics for actual M results - Sardinia (2012)*

Variable	Obs	Mean	Std. Dev.	Min	Max
M (5 km)	219	1,43	1,73	0	112,28
M (10 km)	219	1,28	0,95	0	66,82
M (15 km)	219	1,20	0,68	0	52,02
M (20 km)	219	1,16	0,50	0	39,21
M (30 km)	219	1,10	0,23	0	11,13

Source: Compiled by the authors

Table 3 – *Statistics for percentage changes in actual M in Sardinia between 2007 and 2012*

Variable	Obs	Mean	Std. Dev.	Min	Max
M (5 km)	216	5,32	262,86	-100	34195,32
M (10 km)	216	4,45	73,85	-100	7033,24
M (15 km)	216	3,85	52,97	-100	5391,64
M (20 km)	216	3,96	42,37	-100	3685,38
M (30 km)	216	3,66	17,87	-100	1605,94

Source: Compiled by the authors

Table 4 – *Correlation between actual M and approximated M*

	Approximated M (Km)	Approximated M (Minutes)	Δ 2007-2012 for Approximated M (Km)	Δ 2007-2012 for Approximated M (Minutes)
Actual M (5 km)	0,9	0,75	1,00	0,99
Actual M (10 km)	0,84	0,85	0,83	0,83
Actual M (15 km)	0,96	0,96	0,98	0,81
Actual M (20 km)	0,97	0,97	0,96	0,96
Actual M (30 km)	0,89	0,88	0,88	0,86

Source: Compiled by the authors

Table 5 – *20 industries most agglomerated within a 15 km radius in Sardinia in 2007 (only industries with ≥10 plants)*

Industry code	Industry description	Employees	Plants	M (15 km)
257	MANUFACTURE OF CUTLERY, HAND TOOLS AND GENERAL HARDWARE	58,58	38	7,70
143	MANUFACTURE OF KNITTED AND CROCHETED APPAREL	303,38	17	6,58
553	CAMPING GROUNDS, RECREATIONAL VEHICLE PARKS AND TRAILER PARKS	676,97	102	5,18
501	SEA AND COASTAL WATER TRANSPORT	516,28	100	5,05
201	MANUFACTURE OF BASIC CHEMICALS, FERTILIZERS AND NITROGEN COMPOUNDS, PLASTICS AND SYNTHETIC RUBBER IN PRIMARY FORMS	2.059,97	48	4,01
212	MANUFACTURE OF MEDICINAL CHEMICAL AND BOTANICAL PRODUCTS	53,02	16	3,74
81	QUARRYING OF STONE, SAND AND CLAY	1464,57	247	3,49
244	MANUFACTURE OF BASIC PRECIOUS AND OTHER NON-FERROUS METALS; REPROCESSING OF NUCLEAR FUELS	1.794,96	21	3,43
192	MANUFACTURE OF REFINED PETROLEUM PRODUCTS	1.412,37	27	3,03
552	OTHER SHORT TERM ACCOMODATION ACTIVITIES	2.098,36	414	2,97
559	OTHER ACCOMODATION	102,89	23	2,64
109	MANUFACTURE OF PREPARED ANIMAL FEEDS	277,21	35	2,64
551	HOTELS	8.476,83	835	2,62
105	MANUFACTURE OF DAIRY PRODUCTS	2.020,45	177	2,53
162	MANUFACTURE OF PRODUCTS OF WOOD, CORK, STRAW AND PLAITING MATERIALS	4.974,95	1469	2,52
103	PROCESSING AND PRESERVING OF FRUIT AND VEGETABLES	573,58	60	2,46
271	MANUFACTURE OF ELECTRIC MOTORS, GENERATORS, TRANSFORMERS AND ELECTRICITY DISTRIBUTION AND CONTROL APPARATUS	137,18	13	2,44
233	MANUFACTURE OF CLAY BUILDING MATERIALS	397,33	18	2,41
324	MANUFACTURE OF GAMES AND TOYS	53,56	16	2,40
131	SPINNING, WEAVING AND FINISHING OF TEXTILES	219,73	25	2,39

Source: Compiled by the authors

Table 6 – *20 industries most agglomerated within a 15 km radius in Sardinia in 2012 (only industries with ≥10 plants)*

Industry code	Industry description	Employees	Plants	M (15 km)
133	FINISHING OF TEXTILES	10,4	10	26,84
143	MANUFACTURE OF KNITTED AND CROCHETED APPAREL	185,44	13	15,80
271	MANUFACTURE OF ELECTRIC MOTORS, GENERATORS, TRANSFORMERS AND ELECTRICITY DISTRIBUTION AND CONTROL APPARATUS	49,8	14	6,87
257	MANUFACTURE OF CUTLERY, HAND TOOLS AND GENERAL HARDWARE	53,55	34	5,28
283	MANUFACTURE OF AGRICULTURAL AND FORESTRY MACHINERY	32,91	11	5,04
501	SEA AND COASTAL WATER TRANSPORT	593,84	114	4,76
132	WEAVING OF TEXTILES	248,12	30	4,66
553	CAMPING GROUNDS, RECREATIONAL VEHICLE PARKS AND TRAILER PARKS	502,42	85	4,53
192	MANUFACTURE OF REFINED PETROLEUM PRODUCTS	1.448,77	31	4,20
201	MANUFACTURE OF BASIC CHEMICALS, FERTILIZERS AND NITROGEN COMPOUNDS, PLASTICS AND SYNTHETIC RUBBER IN PRIMARY FORMS	1.288,52	34	4,03
151	TANNING AND DRESSING OF LEATHER; MANUFACTURE OF LUGGAGE; HANDBAGS; SADDLERY AND HARNESS; DRESSING AND DYEING OF FUR	37,46	27	3,86
105	MANUFACTURE OF DAIRY PRODUCTS	1.673,86	161	3,26
642	ACTIVITIES OF HOLDING COMPANIES	2,56	20	3,15
551	HOTELS	7.000,48	744	3,05
301	BUILDING OF SHIPS AND BOATS	139,27	58	3,02
162	MANUFACTURE OF PRODUCTS OF WOOD, CORK, STRAW AND PLAITING MATERIALS	3.431,71	1088	2,97
552	OTHER SHORT TERM ACCOMODATION ACTIVITIES	1.710,03	644	2,95
103	PROCESSING AND PRESERVING OF FRUIT AND VEGETABLES	162,44	31	2,93
109	MANUFACTURE OF PREPARED ANIMAL FEEDS	143,8	23	2,88
235	MANUFACTURE OF CEMENT, LIME AND PLASTER	222,23	10	2,87

Source: Compiled by the authors

Table 7 *– 20 industries most dispersed within a 15 km radius in Sardinia in 2007 (only industries with ≥10 plants)*

Industry code	Industry description	Employees	Plants	M (15 km)
279	MANUFACTURE OF OTHER ELECTRICAL EQUIPMENT	201,74	19	0,28
502	INLAND WATER TRANSPORT	93,5	12	0,40
263	MANUFACTURE OF COMMUNICATION EQUIPMENT	128,65	17	0,61
429	CONSTRUCTION OF OTHER CIVIL ENGINGEERING PROJECTS	951,51	169	0,72
203	MANUFACTURE OF PAINTS, VARNISHES AND SIMILAR COATINGS, PRINTING INK AND MASTICS	162,4	35	0,75
853	SECONDARY EDUCATION	118,7	25	0,75
619	OTHER TELECOMMUNICATIONS ACTIVITIES	160,72	45	0,76
611	WIRED TELECOMMUNICATIONS ACTIVITIES	1.816,50	58	0,79
782	TEMPORARY EMPLOYMENT AGENCY ACTIVITIES	2.639,45	65	0,80
812	CLEANING ACTIVITIES	9.116,06	1070	0,81
205	MANUFACTURE OF OTHER CHEMICAL PRODUCTS N.E.C.	93,1	19	0,82
381	WASTE COLLECTION	2.927,22	133	0,82
171	MANUFACTURE OF PULP, PAPER AND PAPERBOARD	147,33	10	0,85
562	EVENT CATERING AND OTHER FOOD SERVICE ACTIVITIES	1.695,02	240	0,86
493	OTHER PASSENGER LAND TRANSPORT	3.886,69	672	0,89
931	SPORTS ACTIVITIES	568,56	257	0,89
641	MONETARY INTERMEDIATION	6.437,98	743	0,91
531	POSTAL ACTIVITIES	4.057,42	425	0,91
432	ELECTRICAL, PLUMBING AND OTHER CONSTRUCTION INSTALLATION ACTIVITIES	11.800,81	3578	0,93
742	PHOTOGRAPHIC ACTIVITIES	550,28	370	0,94

Source: Compiled by the authors

Table 8 *– 20 industries most dispersed within a 15 km radius in Sardinia in 2012 (only industries with ≥10 plants)*

Industry code	Industry description	Employees	Plants	M (15 km)
204	MANUFACTURE OF SOAP AND DETERGENTS, CLEANING AND POLISHING PREPARATIONS, PERFUMES AND TOILET PREPARATIONS	47,82	13	0,02
559	OTHER ACCOMODATION	27,73	15	0,23
279	MANUFACTURE OF OTHER ELECTRICAL EQUIPMENT	100,38	18	0,41
203	MANUFACTURE OF PAINTS, VARNISHES AND SIMILAR COATINGS, PRINTING INK AND MASTICS	105,45	23	0,48
239	MANUFACTURE OF OTHER NON-METALLIC MINERAL PRODUCTS N.E.C.	67,87	23	0,50
205	MANUFACTURE OF OTHER CHEMICAL PRODUCTS N.E.C.	88,99	17	0,54
491	PASSENGER RAIL TRANSPORT, INTERURBAN	495,16	11	0,56
221	MANUFACTURE OF RUBBER PRODUCTS	262,57	24	0,65
263	MANUFACTURE OF COMMUNICATION EQUIPMENT	61,08	10	0,73
611	WIRED TELECOMMUNICATIONS ACTIVITIES	1.723,74	41	0,74
381	WASTE COLLECTION	3.154,62	143	0,79
243	CASTING OF SEMI-FINISHED STEEL PRODUCTS	70,05	20	0,84
352	MANUFACTURE OF GAS; DISTRIBUTION OF GASEOUS FUELS THROUGH MAINS	68,08	14	0,84
493	OTHER PASSENGER LAND TRANSPORT	4.956,76	727	0,86
411	PROJECT MANAGEMENT ACTIVITIES RELATED TO CONSTRUCTION	27,03	34	0,87
856	EDUCATIONAL SUPPORT ACTIVITIES	30,84	24	0,88
332	INSTALLATION OF INDUSTRIAL MACHINERY AND EQUIPMENT	575,42	157	0,90
871	RESIDENTIAL NURSING CARE FACILITIES	708,41	33	0,91
152	MANUFACTURE OF FOOTWEAR	56,61	23	0,92
641	MONETARY INTERMEDIATION	5.354,60	633	0,93

Table 9 – *20 industries with the largest % increase in agglomeration in Sardinia between 2007 and 2012 (only industries with ≥10 plants in 2007)*

Industry code	Industry description	Δ Employees	Δ Plants	Δ M (15 km)
132	WEAVING OF TEXTILES	-67,55	-18,92	206,31
271	MANUFACTURE OF ELECTRIC MOTORS, GENERATORS, TRANSFORMERS AND ELECTRICITY DISTRIBUTION AND CONTROL APPARATUS	-63,70	7,69	182,17
143	MANUFACTURE OF KNITTED AND CROCHETED APPAREL	-38,88	-23,53	140,13
283	MANUFACTURE OF AGRICULTURAL AND FORESTRY MACHINERY	-68,99	-47,62	134,02
772	RENTING AND LEASING OF PERSONAL AND HOUSEHOLD GOODS	-27,60	-17,17	109,21
151	TANNING AND DRESSING OF LEATHER; MANUFACTURE OF LUGGAGE; HANDBAGS; SADDLERY AND HARNESS; DRESSING AND DYEING OF FUR	-56,28	-40,00	101,66
429	CONSTRUCTION OF OTHER CIVIL ENGINGEERING PROJECTS	2,11	-3,55	94,10
274	MANUFACTURE OF ELECTRIC LIGHTING EQUIPMENT	-8,68	-26,32	78,22
619	OTHER TELECOMMUNICATIONS ACTIVITIES	45,74	73,33	71,53
329	OTHER MANUFACTURING N.E.C.	-0,09	20,41	70,45
782	TEMPORARY EMPLOYMENT AGENCY ACTIVITIES	0,38	-21,54	66,70
421	CONSTRUCTION OF ROADS AND RAILWAYS	-8,11	-1,74	63,82
89	MINING AND QUARRYING N.E.C.	-12,18	-14,29	56,34
803	INVESTIGATION ACTIVITIES	140,78	65,00	51,86
301	BUILDING OF SHIPS AND BOATS	-51,54	-25,64	48,86
235	MANUFACTURE OF CEMENT, LIME AND PLASTER	-25,08	0,00	47,29
262	MANUFACTURE OF COMPUTERS AND PERIPHERAL EQUIPMENT	-61,62	-52,83	45,33
279	MANUFACTURE OF OTHER ELECTRICAL EQUIPMENT	-50,24	-5,26	44,77
101	PROCESSING AND PRESERVING OF MEAT	-17,83	-4,35	41,13
873	RESIDENTIAL CARE ACTIVITIES FOR THE ELDERLY AND DISABLED	12,10	-21,39	40,44

Source: Compiled by the authors

Table 10 – *20 industries with the largest % decrease in agglomeration in Sardinia between 2007 and 2012 (only industries with ≥10 plants in 2007)*

Industry code	Industry description	Δ Employees	Δ Plants	Δ M (15 km)
559	OTHER ACCOMODATION	-73,05	-34,78	-91,34
491	PASSENGER RAIL TRANSPORT, INTERURBAN	-75,33	-87,50	-73,08
239	MANUFACTURE OF OTHER NON-METALLIC MINERAL PRODUCTS N.E.C.	-73,44	-20,69	-72,11
411	PROJECT MANAGEMENT ACTIVITIES RELATED TO CONSTRUCTION	-91,47	-65,66	-58,70
131	SPINNING, WEAVING AND FINISHING OF TEXTILES	-66,60	-52,00	-45,97
221	MANUFACTURE OF RUBBER PRODUCTS	15,72	20,00	-43,90
871	RESIDENTIAL NURSING CARE FACILITIES	-2,32	-51,47	-42,30
172	MANUFACTURE OF CORRUGATED PAPER AND PAPERBOARD AND OF CONTAINERS OF PAPER AND PAPERBOARD	-14,27	-18,18	-39,72
203	MANUFACTURE OF PAINTS, VARNISHES AND SIMILAR COATINGS, PRINTING INK AND MASTICS	-35,07	-34,29	-36,17
205	MANUFACTURE OF OTHER CHEMICAL PRODUCTS N.E.C.	-4,41	-10,53	-34,76
532	COURIER ACTIVITIES	49,36	51,43	-33,62
233	MANUFACTURE OF CLAY BUILDING MATERIALS	1,04	27,78	-33,23
257	MANUFACTURE OF CUTLERY, HAND TOOLS AND GENERAL HARDWARE	-8,59	-10,53	-31,48
139	MANUFACTURE OF OTHER TEXTILES	-27,39	-28,08	-29,46
352	MANUFACTURE OF GAS; DISTRIBUTION OF GASEOUS FUELS THROUGH MAINS	-41,73	-17,65	-28,34
81	QUARRYING OF STONE, SAND AND CLAY	-39,78	-38,46	-26,13
243	CASTING OF SEMI-FINISHED STEEL PRODUCTS	-57,23	-42,86	-24,79
282	MANUFACTURE OF OTHER GENERAL-PURPOSE MACHINERY	-59,79	-25,00	-24,75
152	MANUFACTURE OF FOOTWEAR	-21,63	15,00	-19,21
799	OTHER RESERVATION SERVICE AND RELATED ACTIVITIES	21,53	33,67	-19,21

Source: Compiled by the authors

Ultimi Contributi di Ricerca CRENoS

I Paper sono disponibili in: http://www.crenos.unica.it

22/05 *Luca Scaffidi Domianello, Giampiero M. Gallo, Edoardo Otranto,* "Smooth and Abrupt Dynamics in Financial Volatility: the MS-MEM-MIDAS"

22/04 *Claudio Detotto, Riccardo Marselli, Bryan C. McCannon, Marco Vannini,* "Experts and Arbitration Outcomes: Insights from Public Procurement Contract Disputes"

22/03 *Gerardo Marletto, Cécile Sillig,* "Global influence and national diversity in socio-technical transitions: a sectoral taxonomy"

22/02 *Federico Aresu, Emanuela Marrocu, Raffaele Paci,* "Public capital and institutions' quality in the Italian regions"

22/01 *Gabriele Cardullo, Maurizio Conti, Andrea Ricci, Sergio Scicchitano, Giovanni Sulis,* "On the Emergence of Cooperative Industrial and Labor Relations"

21/11 *Gianfranco Atzeni, Luca G. Deidda, Marco Delogu, Dimitri Paolini,* "Effects of migration with endogenous labor supply and heterogeneous skills"

21/10 *Alberto Tidu, Stefano Usai, Frederick Guy,* "Agglomeration in manufacturing and services: an experimental application of a distance-based measure to Sardinia"

21/09 *Andrea Caria, Fabio Cerina, Marco Nieddu,* "Political Selection and Monetary Incentives in Local Parliamentary Systems"

21/08 *Giorgio Garau, Alessio Tola, Maria Veronica Camerada, Salvatore Lampreu, Silvia Carrus,* "Economic and social polarization dynamics in the EU"

21/07 *Emanuela Marrocu, Raffaele Paci, Stefano Usai,* "Direct and indirect effects of universities on European regional productivity"

21/06 *Pasqualina Arca, Gianfranco Atzeni, Luca Deidda,* "The Signalling Role of Trade Credit on Loan Contracts: Evidence from a Counterfactual Analysis"

21/05 *Grazia Sveva Ascione, Laura Ciucci, Claudio Detotto, Valerio Sterzi,* Do universities look like patent trolls? An Empirical Study of University Patent Infringement Litigation in the United States

21/04 *Michele Battisti, Massimo Del Gatto, Antonio Francesco Gravina, Christopher F. Parmeter,* "Robots versus labor skills: a complementarity / substitutability analysis"

21/03 *William Addessi, Marco Delogu,* "Infrastructure Accumulation in Developing Countries: the Role of the Informal Sector"

21/02 *Luca De Benedictis, Vania Licio, Anna Maria Pinna,* "From the historical Roman road network to modern infrastructure in Italy"

21/01 *Silvia Balia, Rinaldo Brau, Marco G. Nieddu,* "Depowering Risk: Vehicle Power Restriction and Teen Driver Accidents in Italy"

20/08 *Giampiero M. Gallo, Demetrio Lacava, Edoardo Otranto,* "On Classifying the Effects of Policy Announcements on Volatility"

20/07 *Luc Bauwens, Edoardo Otranto,* "Modelling Realized Covariance Matrices: a Class of Hadamard Exponential Models"

20/06 *Demetrio Lacava, Giampiero M. Gallo, Edoardo Otranto,* "Measuring the Effects of Unconventional Policies on Stock Market Volatility"

20/05 *Gianfranco Atzeni, Luca G. Deidda, Marco Delogu, Dimitri Paolini,* "Drop-out decisions in a cohort of Italian university students"

20/04 *Emanuela Marrocu, Raffele Paci, David Rigby, Stefano Usai,* "Smart Specialization Strategy: any relatedness between theory and practice?

20/03 *Giorgio Garau, Stefano Deriu,* "Total Factor Productivity and Relative Prices: the case of Italy"

20/02 *Fabio Cerina, Alessio Moro, Michelle Rendall,* "A Note on Employment and Wage Polarization in the U.S."

www.crenos.unica.it

ISBN 9788868514389